587 Background

[Most of this should be familiar ...]

(AS= Alon-Spencer, JLR= Janson-Łuczak-Ruciński, RG= Bollobás' *Random Graphs*, MR= Molloy-Reed.)

Notation

$\mathbb{N} = \{0, 1, \ldots\}$, $\mathbb{P} = \{1, 2, \ldots\}$, $[n] = \{1, \ldots, n\}$

$(n)_k = n(n-1)\cdots(n-k+1)$ $(= k!\binom{n}{k})$

$2^X = \{\text{subsets of } X\}$

$\binom{X}{k} = \{k\text{-subsets of } X\}$

For f, g functions (of whatever, and with limits as that whatever tends to ... something):

$f = O(g)$ means $|f| \leq C|g|$ with C constant;

$f = \Omega(g)$ means $|f| \geq C|g|$ with C a positive constant;

$f = \Theta(g)$ (also written $f \asymp g$) means $f = O(g)$ *and* $f = \Omega(g)$;

$f = o(g)$ (or $f \ll g$) means $f/g \to 0$, $f = \omega(g)$ (or $f \gg g$) means $|f/g| \to \infty$;

$f \sim g$ means $f/g \to 1$;

and $f \approx g$ means nothing precise.

Warning: we'll often (as above) use "hidden parameters," i.e. underlying parameters that we suppress in our notation. This may be slightly confusing initially, but will eventually make our lives considerably easier.

Probability

I won't try to do this for real here, but a few items of particular interest for us are mentioned below. See Ross's *Introduction to Probability Models* for a nice review of basics (if you need them). Beyond these, Ross and Feller are good, clear introductions to probability. (Durrett is also nice, but significantly harder.) Actually, the first chapters of RG and JLR give enough background for most of what we'll do, and there's also a little in MR.

Hypergraphs

For us a *hypergraph* is a collection $\mathcal{H}$ (possibly with repeats) of subsets (called *edges*) of some finite (*vertex*) set V. (Note this notation differs slightly from that in AS, p.9.) One says $\mathcal{H}$ is *r-uniform* if each of its edges has size r; *d-regular* if each of its vertices has degree d (i.e. belongs to exactly d edges); and *simple* (or *nearly-disjoint*) if $|A \cap B| \leq 1$ whenever A, B are distinct edges (a.k.a. any two vertices share at most one edge).

Random variables

Most of our r.v.'s X are *discrete* (that is, they take values in some finite or countable set), so are specified by the probabilities $\mathbb{P}(X = x)$. The mean of a $\mathbb{R}$-valued r.v. X is μ_X or $\mathbb{E}X$, and the variance is σ_X^2 or $\mathrm{Var}(X)$.

Distributions we'll tend to see:

$X \sim \mathrm{Ber}(p)$ means X is *Bernoulli* with parameter p; that is, $X \in \{0,1\}$ with $\mathbb{P}(X=1)=p$. If A is an event then its *indicator* is the Bernoulli r.v. $\mathbb{1}_A$ given by

$$\mathbb{1}_A = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise.} \end{cases}$$

$X \sim \mathrm{Bin}(n,p)$ means X is *binomial* with parameters n,p; i.e. $X = \sum_{i=1}^n X_i$, with the X_i's independent $\mathrm{Ber}(p)$ r.v.s. So X is the number of successes in n independent trials, each with probability of success p, and

$$\mathbb{P}(X=k) = \tbinom{n}{k} p^k (1-p)^{n-k} \; 0 \le k \le n.$$

$X \sim \mathrm{Geom}(p)$ means X is *geometric* with parameter p; that is, $\mathbb{P}(X=i) = (1-p)^{i-1}p$. So we may regard X as the number of trials until success if the trials are independent and each succeeds with probability p.

$X \sim \mathrm{Po}(\mu)$ means X is *Poisson* with mean μ; that is,

$$\mathbb{P}(X=k) = e^{-\mu}\mu^k/k! \text{ for } k \in \mathbb{N}.$$

[Little exercises: (a) find the means and variances of $X \sim \mathrm{Bin}(n,p)$ and $Y \sim \mathrm{Geom}(p)$, and the variance of $Z \sim \mathrm{Po}(\mu)$; (b) what can you say about $X+Y$ if X and Y are independent with $X \sim \mathrm{Po}(\mu)$ and $Y \sim \mathrm{Po}(\lambda)$?]

$X \sim \mathrm{N}(0,1)$ means X is standard normal, i.e. has density function

$$\phi(t) = \tfrac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

and (therefore) distribution function

$$\mathbb{P}(X < x) = \Phi(x) := \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt.$$

Then $\mu_X = 0$, $\sigma_X^2 = 1$, and one can show that as $x \to \infty$,

$$(\mathbb{P}(X \ge x) =)\; 1 - \Phi(x) \sim x^{-1}\phi(x) \; (= \tfrac{1}{x\sqrt{2\pi}} e^{-x^2/2}).$$

More generally, $X \sim \mathrm{N}(\mu,\sigma)$ has density function $\frac{1}{\sqrt{2\pi}} e^{-(t-\mu)^2/(2\sigma^2)}$, mean μ, and variance σ^2.

And we *might* also see: $X \sim \mathrm{Exp}(\lambda)$ means X is *exponential* with parameter λ ($0 \le \lambda < \infty$); that is,

$$\mathbb{P}(X > x) = e^{-\lambda x} \; x \ge 0.$$

Estimates

Stirling's formula:

$$n! \sim \sqrt{2\pi n}\,(n/e)^n.$$

Even better estimates exist, but we won't need them, and *usually* won't even care about the $\sqrt{2\pi n}$.

Approximation by powers of e

For *any* $x \in \mathbb{R}$, $1+x < e^x$. This doesn't give away much if x is small, and in such cases replacing $1+x$ by the easier-to-work-with e^x should be pretty much a reflex. Sometimes we also need a reverse inequality (e.g. to test the accuracy of $(1+x)^n \approx e^{xn}$); for this you can use something like $1+x > e^{x-x^2}$, valid at least for $|x|$ slightly small (exercise to see more precisely where it's okay).

Binomial coefficients

We use these frequently, so you should get comfortable with them; if you're not already, it might help to go through the easy verifications of the following statements.

For any n, k,

$$\left(\tfrac{n}{k}\right)^k \le \tbinom{n}{k} \le \tfrac{n^k}{k!} \le \left(\tfrac{en}{k}\right)^k.$$

If $k \ll n^{2/3}$, then

$$\tbinom{n}{k} \sim \tfrac{n^k}{k!} e^{-k^2/(2n)};$$

in particular, if $k \ll \sqrt{n}$ then

$$\tbinom{n}{k} \sim \tfrac{n^k}{k!}.$$

If $k = \alpha n$ with $\alpha > 0$ fixed, then $\binom{n}{k} \approx 2^{H(\alpha)n}$, where $H(\alpha) = -\alpha \log_2 \alpha - (1-\alpha)\log_2(1-\alpha)$ (the *binary entropy* of α). More precisely (and usually more precisely than we need),

$$\tbinom{n}{k} \sim [2\pi\alpha(1-\alpha)n]^{-1/2}\, 2^{H(\alpha)n}. \tag{1}$$

If k is only assumed to be *close to* αn, then we get approximations of this estimate; for instance, $k \sim \alpha n$ gives $\log \binom{n}{k} \sim H(\alpha)n$, but *not* (1). (Note "$\log f \sim \log g$" is weaker than "$f \sim g$.")

Large deviations

For independent r.v.'s these are treated in great detail in Appendix A of AS. Chapter 2 of JLR is also good (less detailed but more general). We mention just a prototypical example: for $X \sim \mathrm{Bin}(n, 1/2)$ and $k > 0$,

$$\mathbb{P}(X \ge n/2 + k) \ (= 2^{-n} \textstyle\sum\{\binom{n}{i} : i > n/2 + k\}) \ < e^{-2k^2/n}.$$

(Tiny exercise: this is equivalent to AS, Theorem A.1.)

Inclusion-exclusion

For events $A_1, \ldots, A_n$, $A_I := \cap_{i \in I} A_i$ and $S_k := \sum\{\mathbb{P}(A_I) : I \in \binom{[n]}{k}\}$,

$$\mathbb{P}(\cap_{i\in[n]} \bar{A}_i) = \textstyle\sum\{(-1)^{|I|}\mathbb{P}(A_I) : I \subseteq [n]\} = \sum_{k=0}^{n} (-1)^k S_k$$

and, more generally (K. Jordán's formula),

$$\mathbb{P}(\text{exactly } r \text{ of the } A_i \text{ occur}) = \textstyle\sum_{k=r}^{n} (-1)^{k+r}\binom{k}{r} S_k.$$

Moreover the partial sums here alternately over- and underestimate the left hand side; succinctly:

$$(-1)^{l-r} \textstyle\sum_{k=r}^{l} (-1)^{k+r}\binom{k}{r} S_k \ge 0$$

for each $l \ge r$. For $r = 0$, these are the *Bonferroni Inequalities*; I don't know if the general version has a name.

Convergence

A sequence $\{X_n\}$ of r.v.'s *converges in distribution* to X, if $\mathbb{P}(X_n \le x) \to \mathbb{P}(X \le x) =: F(x)$ for each point of continuity x of F. This is written $X_n \xrightarrow{\mathrm{d}} X$. We also write, e.g., $X_n \xrightarrow{\mathrm{d}} N(0,1)$ when $X_n \xrightarrow{\mathrm{d}} X$ with $X \sim N(0,1)$. Don't worry too much about the definition: we will almost always (or always) have either

(i) X is normal (so every x is a point of continuity); or

(ii) X and the X_n's are discrete, in which case $X_n \xrightarrow{\mathrm{d}} X$ just means that $\mathbb{P}(X_n = x) \to \mathbb{P}(X = x)$ for each fixed x.

[E.g. (exercise): if $X_n \sim \mathrm{Bin}(n, c/n)$ with c fixed, then $X_n \xrightarrow{\mathrm{d}} \mathrm{Po}(c)$.]

Similarly (*not* an exercise): if $X \sim \mathrm{Bin}(n, p)$ with $np(1-p) \to \infty$, then $(X - np)/\sqrt{np(1-p)} \xrightarrow{\mathrm{d}} N(0,1)$. (Note the hidden parameter: p is $p(n)$.) This is a special case of both the DeMoivre-Laplace Theorem (see e.g. RG, p.13 for a very complete statement, the last couple lines of which are what we want) and the Central Limit Theorem.